

# Software

---



Data manipulation and visualization

**Caret package.**

Uniform interface to standardize common modelling tasks. In particular for cross-validation

**XGBoost**

Optimized distributed library for a gradient boosting framework

**xgboostExplainer package.**

Tools to interpret the results of XGBoost



# Linear model

---

All the predictors were significant to the response and a significant regression equation was found:

$$F(4, 7651) = 2.517 \times 10^{-4}, p < 2.2 \times 10^{-16}$$

with

$$R^2 = 0.9294.$$